Tim Vigers

3/26/19

Consulting 2

Group Project

*Propensity Scores*

Introduction

Randomized controlled trials (RCTs) ensure by design that treatment groups have similar baseline characteristics, but these characteristics can differ significantly in observational studies (for example through participant self-selection). It is important to account for these differences before making inferences on treatment effect based on observational data, since one cannot compare outcomes directly as with a RCT. One common method for dealing with this issue is regression adjustment. [1] Propensity scores are simply another way to “remove bias due to all observed covariates.” [2]

There are several reasons that propensity scores are used in place of regression adjustment. First, propensity scores allow for separation between the design and analysis phases of the study because “a matched, stratified, or weighted sample can be constructed

without any reference to the outcome.” [1] Also, research has shown that propensity scores are more effective than regression when the outcome under study is rare

However, once a researcher has decided to use a propensity score approach rather than regression adjustment, there are several questions regarding its implementation. First, how should one estimate the propensity score, and how can one

* Essentially a way of making an observational study look a bit like a RCT.
* Assumes “no unmeasured confounders”
  + So does regression for estimating treatment effects
* Most common method is logistic regression (treatment is the outcome and baseline characteristics are predictors)
  + How to assess the propensity score model?
    - No significance testing (too dependent on sample size)
    - Only baseline, as treatment may affect post-baseline measures
* One to one matching is the most common but not the only option
  + Some debate on whether matched sample observations are independent. Austin argues “observed baseline covariates come from the same multivariate distribution,” therefore not independent.
    - Paired t test or McNemar’s
  + What is the best way to 1-1 match?
  + SAS, R, and Stata
* Stratification on score
  + Mutually exclusive subsets
* IPTW
* Covariate adjustment
* Introduction – description of the issues, when they arise, and why it is challenging or interesting. (1 point)
* Methods – description of one or more standard approaches, including relevant literature (e.g. a few papers) and how to carry out the methods (e.g. steps, software). (2.5 points)
* Results – example(s) of carrying out the methods in real situations like the group might encounter. A real or hypothetical data example and analysis can be effective in some cases. (2.5 points)
* Issues, Controversies, alternate approaches – any controversies or advantages/disadvantages to various approaches, alternate approaches. (1 point)

Summary, conclusions, and recommendations – take-away messages, key references. (1 point)

References

1. Austin, P.C., *An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies.* Multivariate Behav Res, 2011. **46**(3): p. 399-424.

2. Rosenbaum, P.R. and D.B. Rubin, *The central role of the propensity score in observational studies for causal effects.* Biometrika, 1983. **70**(1): p. 41-55.